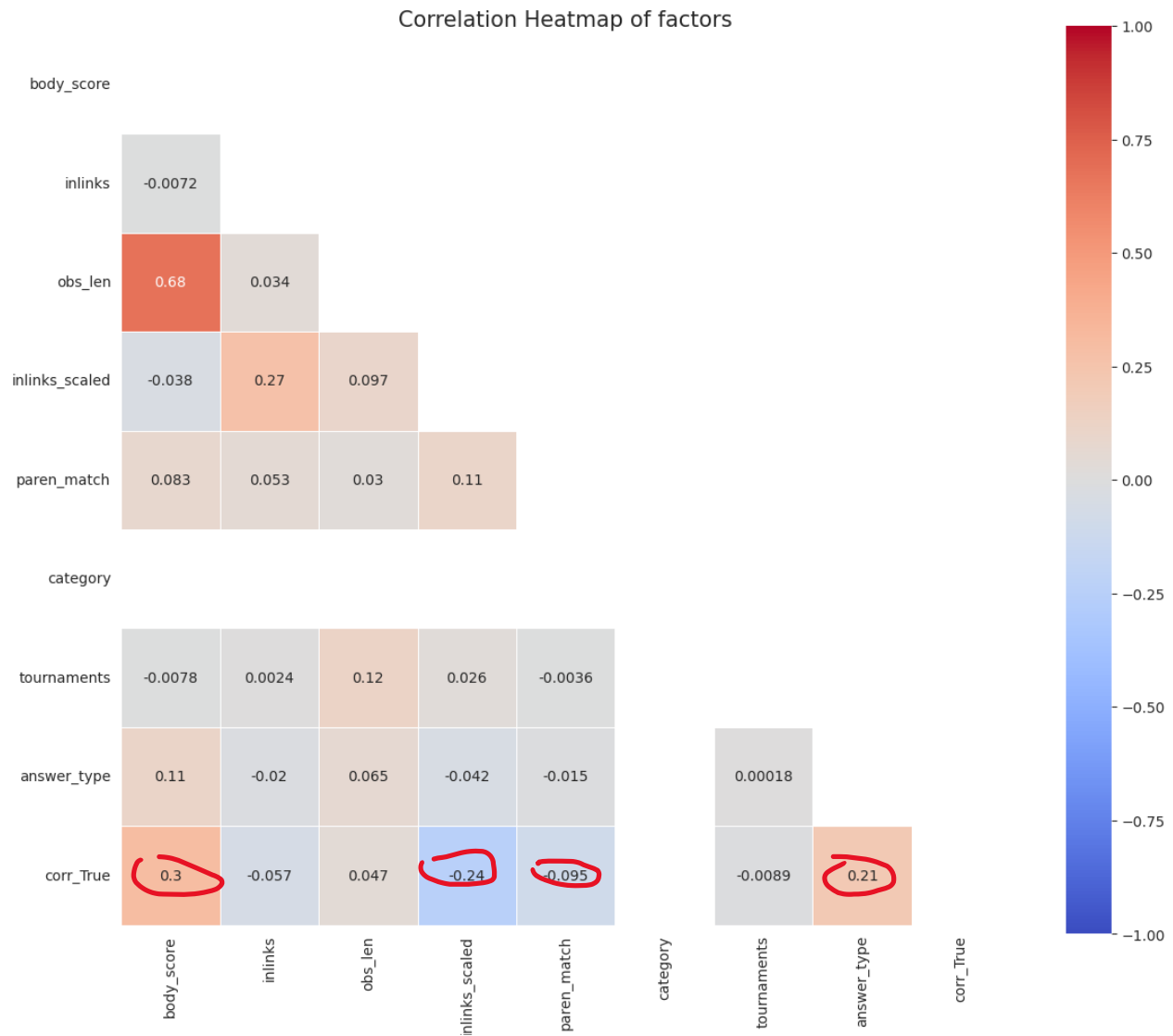


We can see from the correlation matrix that 'body_score', 'inlinks_scaled', 'answer_type', 'paren_match' have high correlations to corr_True, which is the expected output of the model; thus I choose to include them in the training.



We can also see from the result that the accuracy went up from 70 to 80 percent with the confusion matrix below.

